



TITLE:

臨床試験における用量反応性試験 について (種々のモデルの統計的解 析)

AUTHOR(S):

柿爪, 智行

CITATION:

柿爪, 智行. 臨床試験における用量反応性試験について (種々のモデルの統計的解析). 数理解析研究所講究録 2008, 1603: 1-10

ISSUE DATE:

2008-06

URL:

<http://hdl.handle.net/2433/139910>

RIGHT:

臨床試験における用量反応性試験について

アステラス製薬株式会社 データサイエンス部 柿爪 智行 (Tomoyuki Kakizume)
Data Science Dep. Astellas Pharma Inc.

1. はじめに

医薬品開発において、有効性と安全性両方の観点から“至適”な用量を決定することは重要な課題である。医薬品は、一般的に用量が増えるに従って有効性は増加するが安全性は減少する性質をもつ。医薬品を開発する際、用量と臨床での反応の関係（用量反応関係）を知り、安全性が確保できた中でなるべく高い安全性・有効性をもつ用量を“至適”な用量として選択すべきである。

ICH E-4 ガイドライン[1]にも、「医薬品の開発過程には、その医薬品の安全性および有効性の確立に当然あるべき事項として用量－反応の評価のためにデザインされた試験を含めるべき」と記されている。

ここでは、用量反応関係を探索・検証することを目的とした用量反応性試験について、試験の目的・デザイン及び解析手法を[1]および上坂（2007）[4]に従い紹介する。最後に、用量反応性試験を計画・実施する際に注意すべきと個人的に考えている点をまとめる。

2. 用量反応性試験の目的

用量反応性試験の大きな目的として、「用量反応関係の確認」と「検証的試験で用いる用量の選択」が挙げられる。

2.1. 用量反応関係の確認

主に第Ⅱ相試験の初期段階で、用量反応関係の有無の確認および初期的推測を行うことを目的とした用量反応性試験が行われる。ここで得られた情報から、安全かつ有効であると考えられる用量の範囲を推定する。通常、この段階で始めて患者に対して医薬品を投与するため、安全性に配慮した漸増デザインが用いられる。漸増デザインの内容については、3章に示す。

2.2. 検証的試験で用いる用量の選択

安全かつ有効な用量の範囲が推定されたら、その範囲内から複数用量を選択し、それらを用いて検証的試験で用いる推奨用量を選択することを目的とした用量反応性試験を行う。この試験の後、選択された推奨用量および対照薬を用いて検証的試験が行われる。

一般的に、医薬品における用量反応の関係はシグモイド曲線を示す（図1）。そのため、用量反応性試験には通常、前段階までに得られた情報から推定される①有効性が増加し始める用量、②有効性が十分期待できる用量、③これ以上の有効性が期待できない用量が用いられる。また、それら3用量の他に、有効性を正しく評価するための対照としてプラセボを用いることが推奨されている。

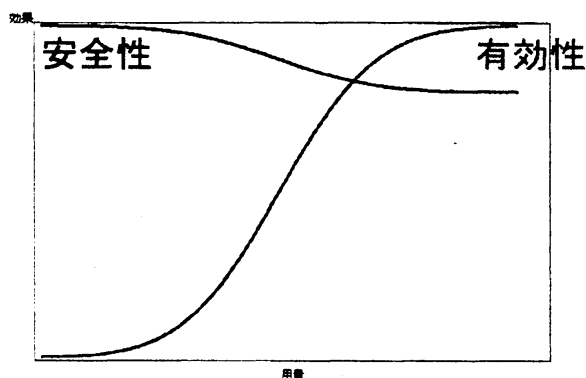


図1. 医薬品における一般的な用量反応関係

検証的試験で用いる推奨用量を選択することを目的とする場合、並行群間デザインが基本となる。ただし、安全性等の問題により、開始時から高用量を投与できない場合には、漸増デザインの利用も考えられる。

3. 用量反応性試験のデザイン

用量反応性試験では、試験の目的だけでなく、医薬品や対象疾患の特徴により、用いるデザインや解析手法が異なる。ここでは、[1]で推奨されているデザインを[4]に基づき紹介する。その他のデザインについては、[4]を参照のこと。

3.1. 固定用量並行群間比較デザイン

複数の固定用量群に対して被験者を無作為に割り付けるデザインである（図2）。このデザインは、比較的単純なデザインであること、また結果の解釈が容易なことから、用量反応試験を行う際、最も標準的なデザインとして利用されている。

各用量の対象母集団に対する平均的な時間反応関係が推定できることや、対象母集団の平均的な用量反応関係に関する情報が得られることがこのデザインの長所として挙げられる。一方、本デザインの短所は、同一患者には固定用量を投与するため、個々の患者の用量反応関係に関する情報は得られないことや、試験開始時から高用量を投与される患者に対する安全性の懸念があげられる。

用量反応関係の確認のみが目的である場合は、ある特定の用量群間での統計的に有意な差を検出する必要はない。しかし、検証的試験で用いる用量を見つけることが目的の場合は、その用量がプラセボに対して統計的に有意であり、かつ臨床的にも意味のある効果を持つことを証明しなければならない。

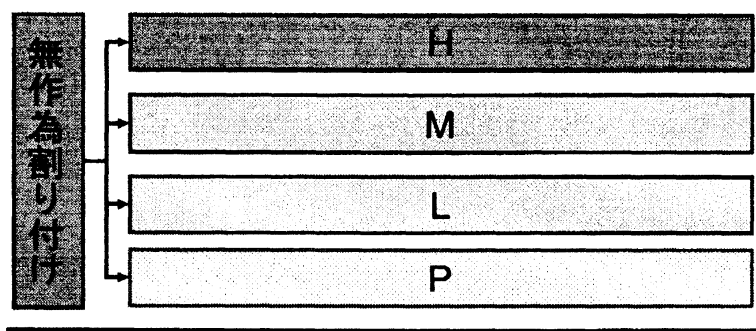


図 2. 固定用量並行群間比較デザイン

3.2. 漸増型並行群間比較デザイン

低用量から投与を開始し、増量計画に基づき割り付けられた用量まで順次増量していくデザインである（図 3）。本デザインでは、割り付けられた用量での効果を比較するため、最終用量を投与する期間は十分長くとる必要がある。

これを用いると、最初から高用量を投与することによる危険を回避しながら、母集団の平均的な用量反応関係を推定することができる。一方、本デザインの短所は、固定用量並行群間比較デザインと同様、個々の患者の用量反応関係に関する情報は得られないことに加え、最終的に評価に用いる用量まで漸増するまでに時間がかかり、試験期間が長くなることが挙げられる。

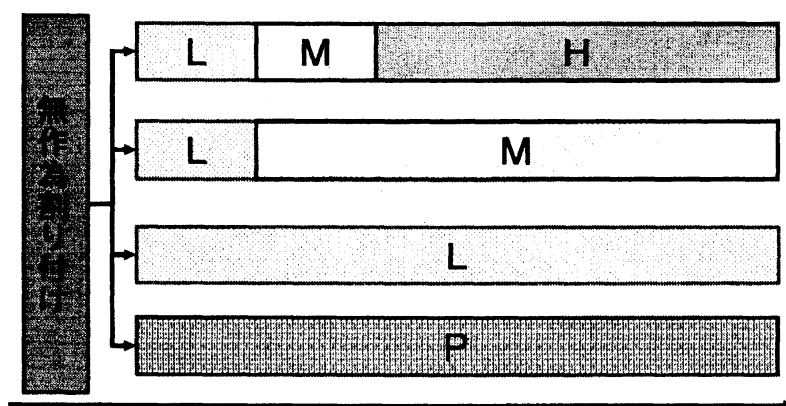


図 3. 漸増型並行群間比較デザイン

3.3. クロスオーバーデザイン

前述の 2 デザインでは基本的に同一患者に固定用量を投与するため、「平均的な」用量反応関係を見ることはできるが、個々の患者の用量反応関係を見ることはできない。医薬品を長期間安全かつ有効に使用するためには適宜用量調整を行う必要があるが、その際の目安として個々の患者の用量反応関係は有用である。平均的な用量反応関係と個々の患者の用量反応関係の両方を見るために有用なデザインが、クロスオーバーデザインである。

クロスオーバーデザインは、同一患者に対して、無作為の順序で複数用量を投与す

るデザインである（図 4）。これは、医薬品の効果が速やかに発現し、かつ中止後には速やかに投与前の状態に戻る場合、有用である。通常は時期効果を考慮して、全員が同じ順序ではなくさまざまな順序で投与するデザインを利用する。また、前に投与された薬剤の効果がその後の効果に影響しないよう、各投与期の間には十分長い休薬期間（wash out 期間）を取る必要がある。

利点は、並行群間デザインに比べて必要な症例数が少なくすむことや、いずれの患者に対しても複数用量が投与されるため、母集団の平均的な用量反応関係だけでなく、個々の患者の用量反応関係も推定することができることである。しかし、各患者に複数用量を投与するため、個々の患者にとって試験期間が極めて長くなることや、途中脱落が多くなると解析が困難になることなどの問題がある。

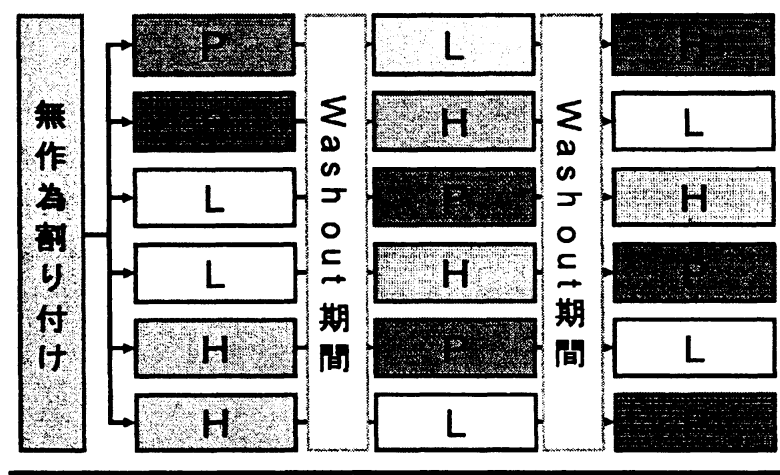


図 4. クロスオーバーデザイン

3.4. 強制的漸増デザイン

強制的漸増デザインは、全ての患者に対して計画された最高用量まで一定期間ごとに順次増量するデザインである。用量の割り付けが無作為ではなく順番であるということ以外は、クロスオーバー用量反応性試験と類似しており、個々の患者の用量反応関係を推定するためのデザインとして適している。

このデザインでは、時間経過に伴い症状が改善する傾向が見られた場合、それが増量したことに起因しているのか、長期間投与したことに起因しているのかを区別することは困難である。特に、副作用の多くは時間依存的な性質を持っているため、このデザインからは副作用についての情報はほとんど得られないという欠点がある。

3.5. 任意漸増デザイン

任意漸増デザインは、十分明確に定義された増量規準に基づき、望ましい反応あるいは望ましくない反応に至るまで投与量を漸増するデザインである（図 5）。安全性に関して患者の差が大きい場合に適用可能であり、初期の用量反応性試験として特に価

定数について、ベクトルを用いて $\mu = (\mu_1, \mu_2, \dots, \mu_K)$, $c = (c_1, c_2, \dots, c_K)$ と表す。このとき、 $\varphi(c, \mu) = \sum_{k=1}^K \mu_k c_k$ を対比という。 $\varphi(c, \mu)$ を用いると、帰無仮説、対立仮説はそれぞれ

$$H_0 : \varphi(c, \mu) = 0,$$

$$H_1 : \varphi(c, \mu) > 0$$

と表すことができる。対比は、さまざまな用量反応関係に対応可能であるため、用量反応関係を確認することが目的である用量反応性試験で用いられることが多い。例えば3用量 ($k=3$) で、用量に関して反応が狭義の単調増加性をもつ ($\mu_1 < \mu_2 < \mu_3$) と仮定するとき、 $c=(-1, 0, 1)$ とすればよい (図 6)。また、用量に関して反応が頭打ちになる ($\mu_1 < \mu_2 = \mu_3$) と仮定する際は、 $c=(-1, 1/2, 1/2)$ とすればよい。

対比を利用した検定において帰無仮説が棄却されたとき、用量反応関係の存在を主張することはできるが、ある特定の用量間における効果の差の存在については知ることとはできない。

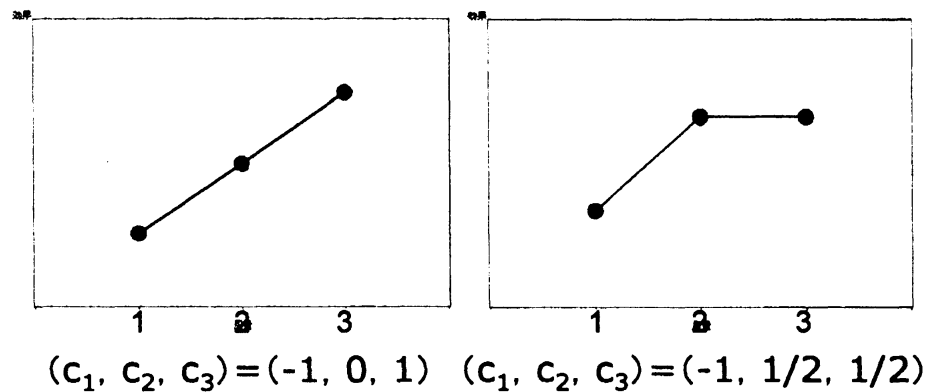


図 6. 対比の例

4.1.1. 正規分布の母平均に対する対比検定

F_k を平均 μ_k 、分散 σ^2 (用量によらず一定) をもつ正規分布 $N(\mu_k, \sigma^2)$ とする。このとき、対比の推定値 $\hat{\varphi}$ とその分散 $v(\hat{\varphi})$ および分散の推定値 $\hat{v}(\hat{\varphi})$ はそれぞれ以下のように表される。

$$\hat{\varphi} = \varphi(c, \bar{Y}) = \sum_{k=1}^K c_k \bar{Y}_k,$$

$$v(\hat{\varphi}) = \sigma^2 \sum_{k=1}^K \frac{c_k^2}{n_k},$$

$$\hat{v}(\hat{\varphi}) = s^2 \sum_{k=1}^K \frac{c_k^2}{n_k}$$

ただし, $\bar{Y}_k = (1/n_k) \sum_{j=1}^{n_k} Y_{kj}$, $s^2 = \sum_{k=1}^K (n_k - 1) s_k^2 / (n - K)$, $s_k^2 = \sum_{j=1}^{n_k} (Y_{kj} - \bar{Y})^2 / (n_k - 1)$ である。

このとき, $\phi := n - K$ とすると, H_0 の下で

$$T_\phi = \frac{\hat{\phi}}{\sqrt{\hat{v}(\hat{\phi})}} \sim t_\phi \text{ 分布 (自由度 } \phi \text{ の } t \text{ 分布)}$$

となるため, t_ϕ 分布の上側 $100\alpha\%$ 点を $t_\phi(\alpha)$ ($0 < \alpha < 1$) とすると, $T_\phi > t_\phi(\alpha)$ ならば有意水準 α で H_0 は棄却され, 用量反応関係はあると判断される。

4.1.2. Cochran-Armitage 検定

エンドポイントが 2 値の場合における, 用量反応性の検定に用いる手法である。

F_k を母数 p_k を持つベルヌーイ分布 $\text{Ber}(p_k)$ とする。このとき, H_0 の下では

$$T_\phi = \frac{\sum_{k=1}^K c_k \bar{Y}_k}{\sqrt{\bar{P}(1-\bar{P}) \sum_{k=1}^K (c_k^2 / n_k)}} \sim N(0,1)$$

が成り立つ。ただし,

$$\bar{Y}_k = \sum_{j=1}^{n_k} Y_{kj}, \quad \bar{P} = \frac{\sum_{k=1}^K \sum_{j=1}^{n_k} Y_{kj}}{\sum_{k=1}^K n_k}$$

である。以上より, $N(0,1)$ の上側 $100\alpha\%$ 点を u_α とすると, $T_\phi > u_\alpha$ ならば有意水準 α で H_0 は棄却される。

Cochran-Armitage 検定は, 用量反応性試験において安全性の検討に用いられることが多いと思われる。その理由としては, 安全性の検討は有害事象の発現率, すなわち有害事象が起こったか起こらなかったかという 2 値データを用いて行われること, また, 安全性の検討を行う際には, 用量反応関係の有無のみチェックし, 特定の群間差があるかどうかという検討は行われることが少ないことが挙げられる。

4.2. 検証的試験で用いる用量の決定が目的の場合

検証的試験で用いる用量の決定が目的の場合, 有効性について, 推奨用量 (候補) のプラセボに対する優越性を示さなければならない。この場合, 各用量とプラセボの対比較を行うため, 多重性の問題が発生する。多重性の問題を解決するため, 用量反応性試験では閉手順を利用した方法が薦められている。

閉手順は, 各仮説に対して重要度を設定し, 重要度の高い順に検定する手法である (図 7)。途中で「有意でない」という結果が出た際は, そこで検定を終了する。閉手順では, 重要度を事前に設定する代り, 各検定では試験全体と等しい有意水準を用い

ることができる。

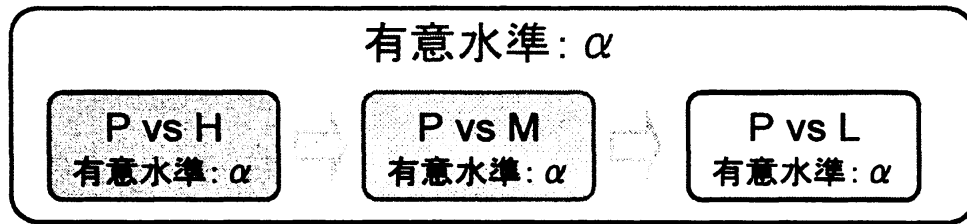


図 7. 閉手順の例 (3 用量 (H>M>L) vs プラセボ (P))

一般的に、医薬品の効果は用量に対して単調が高いほど有効性は高くなると予想できるため、用量反応性試験では高用量から順に検定していくという順序付けは自然であり、結果も解釈がしやすい。

閉手順を利用した手法の 1 つに、Williams 検定がある。Williams 検定は、K 水準の用量について効果の単調増加性を仮定し、高用量から順に検定していく手法である。

各用量群の例数が全て等しく n ($n_1=n_2=\cdots=n_K=n$) の場合を考える。このとき、帰無仮説、対立仮説をそれぞれ(1)のように設定した際の検定は以下の手順で行われる。

手順 1:

各用量群の標本平均 Y_k 、不偏標本分散 V_k ($k=1, 2, \dots, K$)、誤差分散 V_E をそれぞれ算出する。ただし、

$$V_E = \frac{\sum_{k=1}^K (n_k - 1) V_k}{K(n - 1)}$$

とする。

手順 2:

統計量 M_K を算出する。

$$M_K := \max \{T_{2K}, T_{3K}, \dots, T_{KK}\}$$

ただし、

$$T_{iK} := \frac{\sum_{k=i}^K \sum_{j=1}^n Y_{kj}}{n(K-i+1)} \quad (i=2, 3, \dots, K)$$

また、 M_K 、 V_E を用いて、検定統計量 T_K を算出する。

$$T_K := \frac{M_K - \bar{Y}_1}{\sqrt{(2/n)V_E}} \quad (\text{ただし, } \bar{Y}_1 := \frac{1}{n} \sum_{j=1}^n Y_{1j})$$

手順 3:

検定統計量 T_K と有意水準 α に対応する棄却限界値 $\omega(K, \alpha)$ を比較し、

- $T_K < \omega(K, \alpha) \Rightarrow H_0$ を保留し、検定を終了する。
- $T_K \geq \omega(K, \alpha) \Rightarrow$ 「 μ_K は μ_1 より大きい」と判断し、 $K=K-1$ として手順 2、手順 3

を繰り返す。

$\alpha=0.05, 0.025$ に対する棄却限界値 $\omega(K, \alpha)$ については、「統計的多重比較法の基礎」[7]で数表が与えられているので、そちらを参照のこと。

単調増加性を満たしたとき、 T_k は μ_k の最尤推定量となり、Williams 検定は 2 群の対比較を高用量から順に行う閉手順より性能がよいことが知られている。しかし、単調増加性を満たしていないときは、該当する用量のデータのみを用いて対比較を順に行う閉手順の方が性能がよい。そのため、試験を計画する際には、医薬品に対する用量反応性の有無や用量反応曲線の形状を十分検討した上で用いる検定手法を決定すべきである。

5. 用量反応性試験を実施する際の注意点

ここでは、個人的に感じている、用量反応性試験を実施する際の注意点を述べる。

5.1. デザイン

用量反応性試験に関するデザインはさまざまある。その中から状況にあったものを選択する必要があるが、そのときに重要なことは、用量反応性試験の目的を何に設定するかということである。

例えば、処方開始用量を決定することが目的の場合、母集団の「平均的な」用量反応関係に着目することになるが、その際は並行群間比較デザインが有用と考えられる。医薬品開発では処方開始用量の決定が最優先となること（これが決まらなないと処方ができない）、また、デザインの明朗性及び得られたデータの解釈の容易さから、並行群間比較デザインを用いて用量反応試験を実施することが多いと思われる。一方、用量調整の目安を知りたい場合は、患者「個々の」用量反応関係に着目するが、その際はクロスオーバーデザインを用いるのがよいと考えられる。

5.2. 評価変数

用量反応性試験だけでなく、臨床試験を実施する際には、評価変数を何に設定するかということは非常に重要な問題である。変数を選択する際、大事なことは以下の 2 点であると考えられる。

- 有効性の評価変数には、治療目的に直結した治療効果を評価するものを選択すべきである。
- 有効性と安全性が混在しうる評価変数を用いるべきでない。

一般的に用量が増加するに従い、有効性は増加し、安全性は減少する傾向にあるため、これらが混在すると正しい用量反応関係を確認することができない。場合によっては、用量反応関係が山型になってしまう恐れもあるため、とくに注意が必要と考える。

5.3. 事前情報の入手

用量反応関係の確認は主に前期Ⅱ相試験で行われるが、この段階では通常患者に対する情報は殆ど得られておらず、ここで初めて患者に対して医薬品が投与される。そのため、安全性に十分配慮した試験計画を組まなければならない。また、安全性に関しては、第Ⅰ相試験で十分な情報を得ておく必要がある。また、対象疾患の特徴や治療ガイドラインを理解しておくことが必要である。

5.4. 解析手法

対比を用いた解析を用いた場合、実際に得られたデータの傾向と計画当初に予想していた対比が合致しているとき、検定の検出力は最大になることが知られている。また、1種類の対比のみだと、用量反応性はあるが予想していた対比と異なる傾向を示したとき、検定によって用量反応性が認められない可能性がある。この問題の解決策の一つとして、複数対比を設定しておき、それらの同時分布から有意性を判定する手法がさまざまな論文や書籍で紹介されている。

6. おわりに

医薬品の推奨用量や検証的試験で用いる評価項目を決定する上で、用量反応性試験から得られるデータ及び解析結果から得られる情報は非常に重要である。また、患者に対して最初に医薬品が投与される試験という意味でも、用量反応性試験を行う際には慎重な計画を練る必要があると思われる。

<参考文献>

- [1] 厚生省薬務局審査課. 新医薬品の承認に必要な用量－反応関係の検討のための指針. 1994；医薬審第494号.
- [2] 厚生省薬務局. 臨床試験の一般指針. 1998；医薬審第380号.
- [3] 厚生省医薬安全局審査管理課. 臨床試験のための統計的原則. 1998；医薬審第1047号
- [4] 上坂浩之. 医薬開発のための臨床試験の計画と解析, 朝倉書店, 2006.
- [5] 上坂浩之, 丹後俊郎編集. 臨床試験ハンドブック, 朝倉書店, 2006.
- [6] 丹後俊郎, 宮原秀夫編集. 医学統計学ハンドブック, 朝倉書店, 1995.
- [7] 永田靖, 吉田道弘. 統計的多重比較法の基礎, サイエンティスト社, 1997.
- [8] 浜田知久馬. 論文発表のための統計学, 真興交易医書出版部, 1999.